

<https://helda.helsinki.fi>

---

## Electrophysiological correlates of the spatial temporal order judgment task

Simon, Júlia

2020-10

---

Simon , J & Balla , V R 2020 , ' Electrophysiological correlates of the spatial temporal order judgment task ' , Biological Psychology , vol. 156 , 107947 . <https://doi.org/10.1016/j.biopsycho.2020.107947>

---

<http://hdl.handle.net/10138/321456>

<https://doi.org/10.1016/j.biopsycho.2020.107947>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*



# Electrophysiological correlates of the spatial temporal order judgment task

Júlia Simon<sup>a,b,\*</sup>, Viktória Roxána Balla<sup>c</sup>

<sup>a</sup> Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

<sup>b</sup> Department of Cognitive Science, Faculty of Natural Sciences, Budapest University of Technology and Economics, Budapest, Hungary

<sup>c</sup> Cognitive Brain Research Unit, Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

## ARTICLE INFO

### Keywords:

Temporal order judgment  
Auditory perception  
EEG  
ERP  
Temporal discrimination

## ABSTRACT

The study investigated auditory temporal processing on a tens of milliseconds scale that is the interval when two consecutive stimuli are processed either together or as distinct events. Distinctiveness is defined by one's ability to make correct order judgments of the presented sounds and is measured via the spatial temporal order judgement task (TOJ).

The study aimed to identify electrophysiological indices of the TOJ performance. Tone pairs were presented with inter-stimulus intervals (ISI) varying between 25 and 75 ms while EEG was recorded. A pronounced amplitude change in the P2 interval was found between the event-related potential (ERP) of tone pairs having ISI = 55 and 65 ms, but it was a characteristic only of the group having poor behavioral thresholds. With the two groups combined, the amplitude change between these ERPs in the P2 interval showed a medium-size correlation with the behavioral threshold.

## 1. Introduction

Understanding the auditory temporal processing functions of the human cognitive system is crucial to improve differentiation and training of groups displaying deficiencies in linguistic and other domains (Halliday, Tuomainen, & Rosen, 2017). However, some of these functions are still poorly defined and the same putative construct is often measured using different tasks (Fostick & Babkoff, 2013). Furthermore, psychophysical measurements typically reflect the outcome of complex interactions between the putative temporal processing function, task-related strategies, motivation, and other sensory and executive functions (see, e.g., Simon, Takács, Orosz, Berki, & Winkler, 2020). Here, we aim to identify electrophysiological correlates of the auditory spatial temporal order judgment task that is assumed to measure temporal discrimination on a tens of milliseconds scale. Suboptimal temporal discrimination of auditory events of a few tens of milliseconds were proposed to be partly underlying developmental dyslexia (Farmer & Klein, 1995; Fostick, Eshcoly, Shtibelman, Nehemia, & Levi, 2014; Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007), hearing deficits in the elderly over 60 years (Fink, Churan, & Wittmann, 2005; Saija, Başkent, Andringa, & Akyürek, 2019; Szymaszek, Sereda, Pöppel, & Szelag, 2009), and they were also observed in schizophrenics (Stevenson et al.,

2017). Therefore, electrophysiological correlates of this temporal discrimination function may help to specify the role of temporal processing in these deficits as well as provide a bridge between behavioral measures and the underlying neurophysiological mechanisms.

We focused on the spatial temporal order judgment task (sTOJ) as it has been frequently reported as a sensitive indicator of dysfunction (like aphasia or dyslexia) and its training showed transfer effect to linguistic skills, such as phonological awareness (Fostick, Eshcoly et al., 2014; Gaab et al., 2007; Szelag et al., 2014). This task presumably allows the assessment of the duration of a short-term temporal integration window, which has been suggested to be longer in groups at risk of developing language impairment. A longer integration window makes some time-based (e.g. voice onset time) consonant distinctions more difficult (Steinschneider, Fishman, & Arezzo, 2003; Zaehle, Wüstenberg, Meyer, & Jäncke, 2004; Steinschneider et al., 2004; Szelag et al., 2014; Zaehle, Jäncke, & Meyer, 2007). The sTOJ task requires judging the order of two sounds presented to different ears with a variable inter-stimulus interval (ISI) between the sounds. The discrimination threshold (also can be referred as fusion threshold, Cutting, 1976) denotes the interval, below which the order of the two sounds cannot be established on the basis of temporal information alone. Electrophysiological (EEG) measurements have previously been conducted during the sTOJ task. In such studies,

\* Corresponding author at: Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, 1117 Budapest, Magyar tudósok körútja 2, Budapest, Hungary.

E-mail addresses: [simon.julia@ttk.mta.hu](mailto:simon.julia@ttk.mta.hu), [simon.julia@ttk.hu](mailto:simon.julia@ttk.hu), [simon.julia89@gmail.com](mailto:simon.julia89@gmail.com) (S. Júlia).

<https://doi.org/10.1016/j.biopsycho.2020.107947>

Received 9 April 2020; Received in revised form 4 August 2020; Accepted 13 August 2020

Available online 20 August 2020

0301-0511/© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the primary goal was to find correlates of accurate decision at the individual discrimination threshold. It has been shown that higher judgement accuracy is associated with lower activity in the superior temporal cortex (Bernasconi, Grivel, Murray, & Spierer, 2010), stronger pre-stimulus beta activity (Bernasconi, Manuel, Murray, & Spierer, 2011), and increasingly uncorrelated electric brain activity in the two hemispheres (Bernasconi, Grivel, Murray, & Spierer, 2010). However, these results do not automatically refer to processing one versus two auditory events, because correct responses do not necessarily correspond to perceiving two separate sounds. Correct judgments in the spatial version of the TOJ task are less likely to be based on holistic perception than distinguishing the high-low and low-high pattern in the spectral version of the task (Fostick & Babkoff, 2013; Fostick, Wechsler, & Peretz, 2014). Although, one cannot rule out that correct order judgments can be made through utilizing holistic features distinguishing the left-right from the right-left sound pairs (Kanabus, Szélag, Rojek, & Poppel, 2002). As for the studies of Bernasconi et al. (2010a; and 2010b, 2011), the average behavioral threshold was quite low ( $26.15 \pm 5.04$  ms as opposed to the more frequently found 40–50 ms range) and the accuracy was not especially high ( $65.31 \pm 2.91\%$  at the beginning). Therefore, it is possible that the ISI within the tone pairs frequently fell under the discrimination threshold of individuals, who learned subtle cues of discrimination through feedback. Individual thresholds below 30 ms (either in the cited or in our own study) raise the question whether even the spatial TOJ task reflects temporal discrimination based sound processing.

An electrophysiological marker related to the spatial TOJ threshold would thus benefit interpreting the behavioral results of the sTOJ task.

In our previous study, we found an EEG based marker likely indexing the discrimination threshold addressed by sTOJ (Simon, Balla, & Winkler, 2019); the amplitude difference between the ERPs of tone pairs having ISI = 55 ms and ISI = 65 ms was significantly greater than any other amplitude differences of adjacent ISI ERPs. Adjacent pairs where tone pairs differing in one step of ISI. We were looking for nonlinearity in the ERP differences of adjacent ISI pairs as an electrophysiological index of the emerging second auditory event. However, no significant correlation was observed between the electrophysiological measures and behavioral measures of the discrimination threshold. Thus, the EEG based marker could not be anchored in perception. One possible reason for the lack of this correlation is that, in contrast to the behavioral measurement, the sounds were not attended during the EEG measurement. Therefore, in the current study, we introduced an active condition in which participants performed a target detection task unrelated to the sTOJ task, which required them to attend to the sounds. In addition, performing a task could produce more homogeneous neural responses than the passive condition recorded in our previous study. We expected to be able to replicate our previous findings of the ISI effect in the latency range of P2 component, even in the active condition.

Based on our previous results, we hypothesized a P2 reduction at longer ISIs, especially above ISI = 55 ms. The assumption was that if the second tone is processed as a separate auditory event, the appearance of a second onset event (most likely a second N1) would attenuate the positive P2 component compared to the ERPs of tone pairs with short ISIs leading to fused tones.

Besides replicating our previous findings about a non-linear change in the ERP amplitudes as a function of linear increase in ISI, we expected to find correlation between the changes in electrophysiological responses of adjacent ISI ERPs and the behaviorally measured discrimination threshold. At ISI = 25 ms most people hear fused tones and at ISI = 75 ms most people hear two tones, therefore we expected to find correlation where the difference between the ERPs of tone pairs differing in one ISI step is the greatest.

## 2. Methods

### 2.1. Participants

Thirty young adults (20 female) participated in the study. Their mean age was 22.36 years (19–29 range); 13.13% were left handed while 10% reported ambidexterity. None of the participants had a hearing threshold higher than 25 dB HL measured at frequencies 250, 500, 1000, and 2000, or a between-ear difference higher than 10 dB HL. They reported having no psychiatric or neurological conditions.

Prior to the data collection, participants signed an informed consent approved by the United Ethical Review Board of Hungary. They received moderate financial compensation through a student work organization for taking part in the study.

### 2.2. The timeline of the experimental procedure

The session (see also Fig. 1) started with standard audiometry (see the tested frequencies at 2.1) and it was followed by practice for the behavioral Temporal Order Judgment task (TOJ) and three consecutive TOJ threshold measurements (see section 2.3) in order to gain a proper estimate of the individuals discrimination threshold. Participants were then trained to recognize the target pattern of the active EEG-measurement condition that was supposed to direct attention to the dichotic tones as well. This phase of target detection stopped after performance reached 85% based on the last ten target occurrences. This was followed by preparations for the EEG recording (placing the cap and setting the impedances).

The EEG-recording part comprised two conditions; each condition was presented in a separate sequence of six blocks with a minimum one-minute break between blocks (longer, if the participant needed more time). In the “passive condition”, participants attended to a muted self-selected movie with subtitles and they had no task involving the sounds. In the “active condition”, listeners were instructed to pay attention to the sounds and press a response key whenever a target (triplet) was detected. The condition order was counterbalanced across participants resulting in a “passive-start” and an “active-start” group with 15 participants in each group.

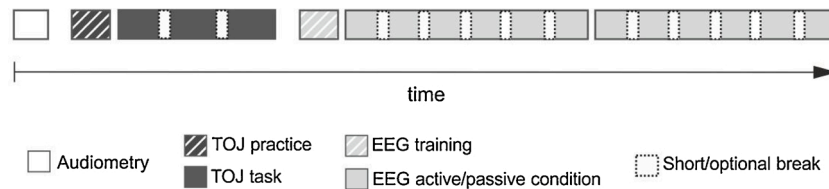
### 2.3. Temporal Order Judgment (TOJ) task

The task consists of a series of order judgments. In the spatial version of the task, the same tone is presented twice in a dichotic manner. Participants are to decide whether they heard a left-to-right or right-to-left propagation of the sounds. In our experiment, the TOJ threshold was measured with a three-down-one-up adaptive algorithm that stopped after eight errors. The initial ISI was 120 ms and the initial step size was 20 ms; the latter was halved after each error until it reached 5 ms. The threshold was calculated as the average of the last six ISIs at which an error was recorded. Participants responded by pressing the ‘1’ or ‘2’ key on a standard IBM PC keyboard with ‘1’ corresponding to the left-to-right sound propagation. Trials (tone pairs to be judged) commenced 600–900 ms after the response to the previous trial. Participants started each threshold measurement whenever they were ready by pressing the SPACE key.

In the practice phase, tone pairs with an ISI of 150 ms were presented six times with feedback after each response. After this accommodation phase, twelve pairs were presented (half of them with ISI = 150 ms and half of them with ISI = 100 ms), and the performance was summarized in a percentage at the end of the block. If the performance was below 85%, this procedure was repeated.

### 2.4. Stimuli

The stimuli for the EEG part were pairs of 800 or 980 Hz pure tones, each tone lasting for 10 ms (with 1–1 ms linear rise and fall times



**Fig. 1.** The structure of the experiment. The behavioral measurements were followed by the EEG recording that had two parts. Participants attended to a silent movie with subtitles in the passive condition, while in the active condition they focused on the tones, because they had to detect a target. The recognition of the target was practiced before the EEG recording started. The passive and active conditions were counterbalanced between participants.

included). Only the 800 Hz tone was used in the behavioral measurement. Two different frequencies were used during the EEG recording in order to reduce the N1 suppression due to adaptation (May & Tiitinen, 2007) and to test whether tone pitch affects the discrimination threshold. Tones were produced by a Juli@ MAYA44 sound card (24-bit, 192 kHz– ESI Audiotechnik GmbH) and delivered via a Sennheiser HD 600 headphone at 68 dB SPL.

The target in the active condition was a sequence of three 10 ms tones in left-right-left order separated by 60 ms silent periods (ISI; see Fig. 2).

Tones were presented in pairs (like in the behavioral task), except for target triplets in the active condition that were inserted between successive pairs of sounds. The time separating the onsets of successive pairs as well as that between a target triplet and a pair was always 850 ms (Fig. 2). The first tone of the pair/triplet was always presented to the left, the second to the right ear. Half of the tone-pairs/triplets (in the active condition, 840/120 pairs/triplets; in the passive condition, 960 pairs with no triplets) were composed of 800-Hz, while the other half from 980 Hz tones. The within-pair ISI varied between 25 and 75 ms in 10 ms steps (six different pairs, in equal number). After creating all different tone patterns (pairs with the 6 different ISIs and the two tone frequencies and triplets with the two tone frequencies), they were randomized with a card-shuffling method applying two constraints: 1) at the beginning of the sequence, six additional tone pairs were presented, each with a different ISI and appearing in random order and 2) each triplet was followed by at least one pair. Altogether, in each condition, 1926 patterns (pairs and triplets, together) were delivered. Both conditions were segmented into blocks of 321 patterns (six blocks of 4.5 min, each) with a minimum 30 s long break between blocks. The break could be extended in the active condition at the participant's request.

## 2.5. EEG recording

EEG was recorded at 1000 Hz and 24 bits resolution with a BrainAmp DC 64-channel amplifier (Brain Products GmbH) and actiCAP active electrodes. During recording, a 0.1 Hz high-pass and a 250 Hz low-pass filters were applied online. The placement of the electrodes followed the

International 10/20 system with three additional electrodes: one placed on the tip of the nose (for the final reference), one below the left, and one lateral to the right eye (serving to assess the composite electrooculogram in a bipolar montage). The FCz electrode served as the common reference during recording.

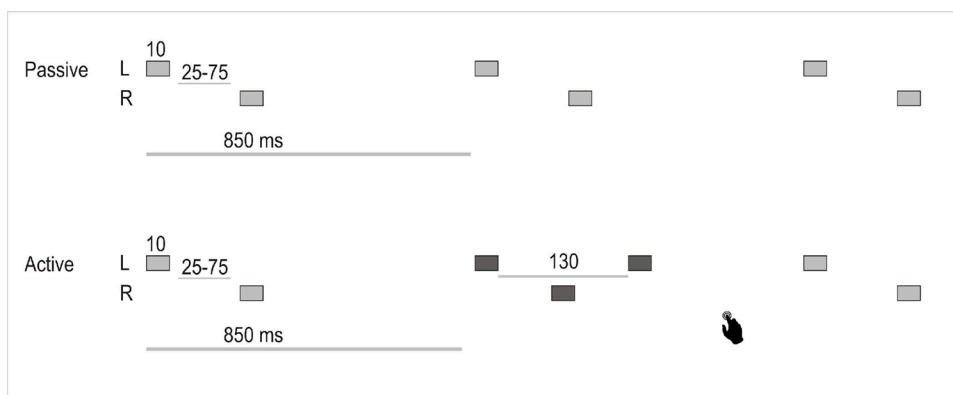
## 2.6. Preprocessing of electrophysiological data

MATLAB R2014a software (MathWorks Inc.) was used for the analyses with custom scripts adopting functions from the EEGLAB (Delorme & Makeig, 2004) and ERPLAB toolboxes (Lopez-Calderon & Luck, 2014). The occasionally malfunctioning P7 recording channel was interpolated, then the signal was re-referenced to the nose electrode and filtered between 1 and 30 Hz with a finite impulse response filter (Kaiser window = 5.6533).

Epochs of 900 ms duration (100 ms pre-, and 800 ms post stimulus) were extracted from the continuous EEG, time locked to the onset of the first sound of the tone pairs. The first six pairs of each stimulus block, the pairs following a triplet in the active condition, as well as those epochs that included >150  $\mu$ V voltage changes during the whole period were rejected from further analysis. Independent Component Analysis (ICA) was applied to the preprocessed data and components related to eye movements were subtracted from the signals based on the recommendations of the ADJUST 1.1.1 software (Mognon, Jovicich, Bruzzone, & Buiatti, 2011).

## 2.7. ERP analysis

Shorter epochs of 500 ms post stimulus interval were cut from the pre-processed epochs. They were baseline corrected by the average voltage in the 100 ms long pre-stimulus period and epochs with >100  $\mu$ V voltage changes occurring during the whole period were rejected from further analysis. An artificial electrode was created by averaging the ERPs from Fz, Cz, FC1, and FC2 for improving the S/N ratio and to maintain compatibility with Simon et al. (2019). Measurements from this electrode were entered into the statistical analyses. ERPs were averaged separately for each condition, pair type (6 ISIs and 2 tone



**Fig. 2.** Stimulus paradigm for the EEG part. Rectangles represent the 10 ms long pure tones. The ISI between the pairs varied between 25 and 75 ms in 10 ms steps. Target triplets appeared only in the active condition. “L” and “R” denote the left and right ears, respectively. In the passive condition, the tone pairs were not actively attended. In the active condition, due to the detection of triplets, attention was directed to the sounds. Three schematic trials are depicted from each condition. The onset asynchrony of these trials was 850 ms.

frequencies), and participant. Although the 800 Hz tones evoked responses with somewhat higher amplitude in the P2 latency range, the data was collapsed across tone frequency, because no significant interaction was found between TONE-FREQUENCY and either the ISI (6 levels) or the CONDITION (passive vs. active) factor (for a comparison figure and the description of the analysis, see Supplementary Material, Section A).

## 2.8. Measurements from the EEG experiment

Since the last 100 ms of the epochs did not contain relevant information, due to more statistical power and better plotting, we only used a 400 ms long post stimulus interval.

In order to find the optimal amplitude measure for testing the effects of the manipulations, we searched for the latency ranges sensitive to the ISI variable, separately in the passive and the active condition. The 400 ms post-stimulus period was divided into 40 consecutive bins of 10 ms duration each, and the average amplitude was calculated for each bin. A repeated measures ANOVA was run on each bin with the within subject factor ISI ( $N = 6$ ) and with a significance criterion level of  $p = .00125$  (taking Bonferroni correction for multiple comparisons into account), separately for the passive- and the active-start group and the passive and the active condition. We found two common intervals sensitive to the within-pair ISI from the above four analyses: 150–170 ms (early window) and 210–230 ms (late window). In these intervals, mean amplitudes and amplitude differences between adjacent ISIs at the artificial electrode were calculated for each participant and condition from these two latency ranges. The aim was to find a non-linear change as a function of ISI, in order to find a potential index of event separation. The physical differences between the tone pairs are expected to generate similar differences between the ERPs differing in one ISI step (adjacent pairs).

In the active condition, hits were defined as a key press no later than 1700 ms (two SOA – stimulus onset asynchrony) after target onset. The number of false alarms was calculated by subtracting hits from all button presses. Hit rates were used to check whether the participants followed instruction.

## 2.9. Statistical analysis

In our previous studies with TOJ (Simon et al., 2020), we found a performance decrement (threshold increase) when measuring TOJ thresholds multiple times in a row. Using a repeated measures ANOVA with the within-subject factor RUN ( $N = 3$ ) we tested this effect on the current data.

Possible differences between the active- and the passive-start group in the mean behavioral TOJ threshold and in the hit rate and the number of false alarms in target detection were tested by independent-samples  $t$ -tests. These analyses checked whether the two groups differed in their TOJ thresholds and whether the order of the two conditions affected performance in the active EEG condition.

Because the test-retest reliability of the sTOJ threshold is not especially high (Fink et al., 2005), for exploring the effects of the TOJ threshold, in addition to testing correlation between the behavioral and EEG measures, a “low-” and a “high-threshold” group were formed using median split of the mean TOJ thresholds. The grouping was then entered into the ANOVA comparing ERP measures (see next paragraph). The two groups were also compared on the hit and false alarm numbers of target detection for testing whether difference in the TOJ threshold affected performance in the active EEG condition.

The effects of the manipulations and groupings were first tested on the mean ERP amplitude measures, separately for the early and the late measurement window by mixed mode ANOVAs with the within-subject factors of CONDITION (active vs. passive) and ISI ( $N = 6$ ) and the between-subject factors of STARTING-CONDITION (active- vs. passive-start) and TOJ-GROUP (low- vs. high-threshold). Post-hoc tests

(pairwise comparisons and F-tests) were Bonferroni corrected.

Pearson's correlations were calculated (because the Kolmogorov-Smirnov tests of normality were not significant) between the ERP amplitude differences (between adjacent ISIs) on the one hand and the average and the maximum TOJ threshold on the other hand (the latter is assumed to be a better indicator of discrimination problems), separately for the two conditions (passive and active) and the two latency ranges (early and late).

Statistical significance was set at  $\alpha = .05$ . The  $\eta^2$  effect size is reported for each significant effect. Whenever the assumption of sphericity was violated based on Mauchly's Test, the degrees of freedom were Greenhouse-Geisser corrected and the corresponding  $\epsilon$  factor is reported. All significant effects are reported.

## 3. Results

### 3.1. Behavioral measures

The mean TOJ threshold was 41.96 ms ( $SD = 26.78$ ), the mean of the maximum thresholds was 57.30 ( $SD = 32.64$ ), with the three successive threshold measurements yielding means of 33.20 ( $SD = 25.12$ ), 44.85 ( $SD = 31.03$ ), and 47.81 ms ( $SD = 34.83$ ). The repeated measures ANOVA revealed a significant threshold deterioration during the three measurements ( $F(2,58) = 5.46$ ,  $MSE = 327.68$ ,  $p = .007$ ,  $\eta^2 = .158$ ). No significant threshold difference was found between the active- and the passive-start group.

The groups formed with median split based on their mean TOJ threshold had thresholds of 20.82 ( $SD = 8.72$ ; low-threshold) and 63.097 ms ( $SD = 21.28$ ; high-threshold).

The mean hit rate in the active condition was 91.38% ( $SD = 7.18\%$ ), suggesting that participants focused on performing the task. There were no significant differences between the active- and the passive-start group in the hit rate ( $t(28) = 1.598$ ,  $p = .121$ ), the number of false alarms ( $t(28) = -1.502$ ,  $p = .144$ ), or the mean TOJ threshold ( $t(28) = -1.497$ ,  $p = .146$ ). There was also no difference in hit rate ( $t(28) = .052$ ,  $p = .959$ ) between the low- and the high-threshold group, but there was a difference in false alarms ( $t(28) = -2.69$ ,  $p = .012$ , the high threshold group producing more incorrect responses).

### 3.2. EEG measures

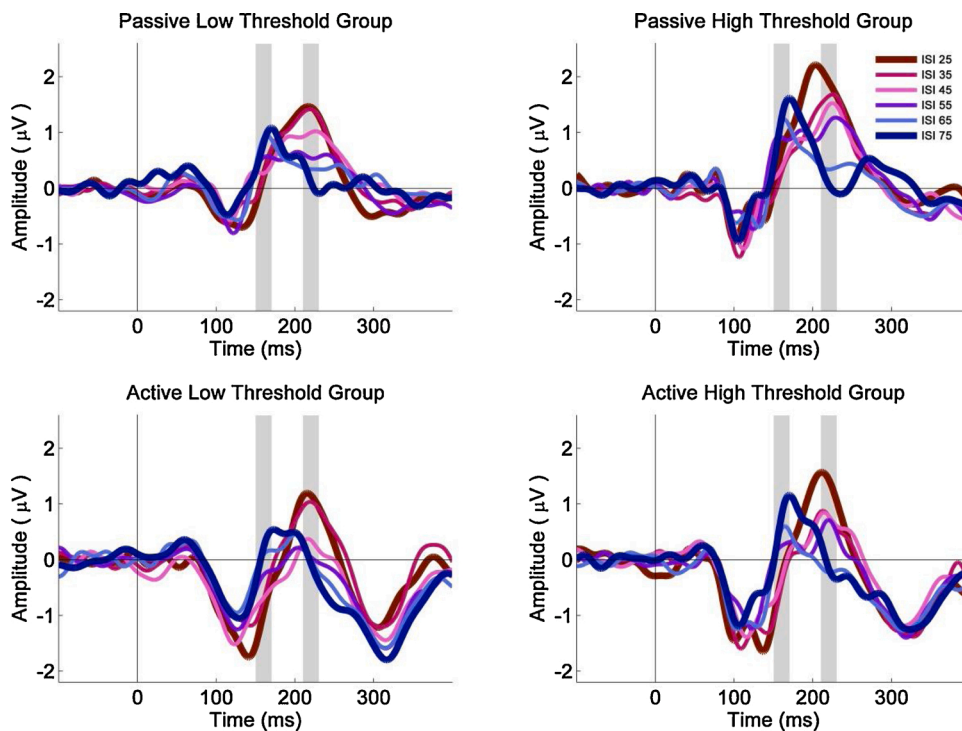
#### 3.2.1. Manipulation and TOJ group effects

The ERPs for the two conditions (passive and active) and the two TOJ-threshold based groups (low- vs. high-threshold) are shown on Fig. 3 with the responses to the different within-pair ISIs overplotted. Visual inspection shows a clear difference in the ERP morphology between short and long ISIs.

**3.2.1.1. Early latency window.** The mixed mode ANOVA of the mean ERP amplitudes yielded significant main effects of CONDITION ( $F(1,26) = 24.29$ ,  $MSE = 2.54$ ,  $p < .001$ ,  $\eta^2 = .415$ ) and ISI ( $F(5,130) = 18.01$ ,  $MSE = .48$ ,  $p < .001$ ,  $\eta^2 = .381$ ). The active condition was associated with more negative amplitudes than the passive condition (Fig. 4A;  $M_{\text{active}} = -.187$  ( $CI95 = [-.657 .283]$ ),  $M_{\text{passive}} = .578$  ( $CI95 = [.172 .984]$ )) and the amplitude generally increased with the ISI. Post-hoc Bonferroni corrected pairwise tests revealed a significant difference only between ISI = 45 and ISI = 55 ( $MD = -.416$ ,  $p = .015$ ; see Fig. 4A showing the effects of both the ISI and the condition).

**3.2.1.2. Late latency window.** The mixed mode ANOVA of the mean ERP amplitudes yielded a significant CONDITION main effect ( $F(1,26) = 7.25$ ,  $MSE = 2.70$ ,  $p = .012$ ,  $\eta^2 = .218$ ; the passive condition amplitudes being more positive than the active condition ones; Fig. 4B), and a significant ISI main effect ( $F(5,130) = 42.82$ ,  $MSE = .45$ ,  $p < .001$ ,  $\eta^2 = .622$ ,  $\epsilon = .668$ ) with a general decrease in amplitude as a function of ISI.





**Fig. 3.** Group mean ( $N = 15$ ) ERP responses to tone pairs by ISI (line color). Upper panels show the responses obtained in the passive, lower panels in the active condition with the low-threshold group on the right and the high-threshold group on the left panels. Data was collapsed across the groups with different starting condition. Amplitude measurement intervals are marked by light grey columns. T0 corresponds to the onset of the first tone. The gray areas indicate the measurement windows (post stimulus 150–170 and post stimulus 210–230) obtained with a series of Bonferroni corrected ANOVAs with the within subject factor ISI.

There was a significant  $ISI \times \text{STARTING-CONDITION}$  interaction ( $F(5,130) = 3.407$ ,  $MSE = .46$ ,  $p = .018$ ,  $\eta^2 = .116$ ,  $\epsilon = .668$ ). No post hoc pairwise group comparison yielded significant difference with the exception of the difference between 45 and 55 ms ISI in the “passive-start” group (see Supplementary Material Section C, Fig. 2). Based on visual inspection, the “passive start” group showed a steeper amplitude decrease as a function of ISI than the active-start group. The  $ISI \times \text{TOJ-GROUP}$  interaction was also significant ( $F(5,130) = 3.34$ ,  $MSE = .45$ ,  $p = .019$ ,  $\eta^2 = .114$ ,  $\epsilon = .668$ ). Again, no post hoc pairwise group comparison was significant. However, the high but not the low threshold group showed a significant amplitude drop between  $ISI = 55$  and  $ISI = 65$  ( $MD = .813$ ,  $p < .001$ ; see Supplementary Material Section C, Fig. 3; the same effect shows up on the depiction of amplitude differences between adjacent ISIs – see, Fig. 4C or Supplementary Fig. 4 for the scalp distributions).

Post hoc repeated measures ANOVA of the amplitude differences between ERPs having adjacent ISIs (collapsed across condition and starting group) was run separately for the two TOJ groups with the within-subject factor  $ISI\text{-PAIR}$  ( $N = 5$ : 35 – 25, 45 – 35, 55 – 45, 65 – 55, 75 – 65). There was no  $ISI\text{-PAIR}$  effect in the ‘low-threshold’ group, but it was significant in the ‘high-threshold’ group ( $F(4,56) = 3.64$ ,  $MSE = .38$ ,  $p = .01$ ,  $\eta^2 = .207$ ). The best polynomial contrast fitting the trend was the cubic function ( $F(1,14) = 12.30$ ,  $MSE = .39$ ,  $p = .003$ ,  $\eta^2 = .468$ ; see also Fig. 4C).

### 3.3. Correlations between electrophysiological and behavioral measures

Fig. 4C suggests that for the low-threshold group, the largest change between adjacent ISIs occurs between 35 and 45 ms, whereas in the high-threshold group, the boundary appears between 55 and 65 ms. This suggests some relationship between the TOJ threshold and the ERP-based indices. The correlation analyses were set up to assess the existence of corroborating evidence.

#### 3.3.1. Early latency window

No significant correlation was found between the amplitude change for any pair of adjacent ISIs and the behavioral TOJ thresholds (mean or

maximum).

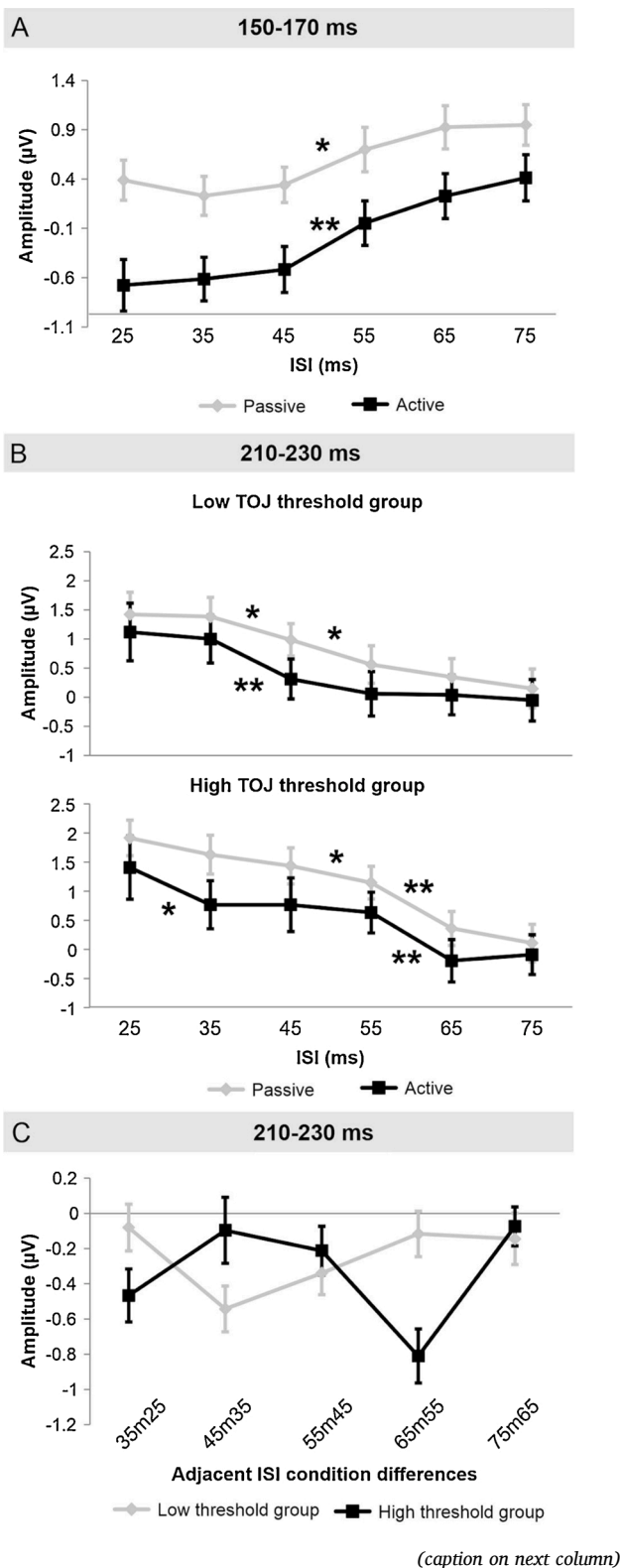
#### 3.3.2. Late latency window

The amplitude difference between 55 and 65 ms ISI significantly correlated with the behavioral threshold data. The correlations in the passive condition were  $\rho(28) = -.444$ ,  $p = .014$  with the mean TOJ threshold and  $\rho(28) = -.424$ ,  $p = .020$  with the maximum TOJ threshold. The correlations in the active condition were  $\rho(28) = -.476$ ,  $p = .008$  with the mean TOJ threshold and  $\rho(28) = -.569$ ,  $p = .001$  with the maximum TOJ threshold. All of the above results hold even after Bonferroni correction. Fig. 5 shows that participants having higher TOJ thresholds (lower temporal resolution) show a more pronounced change between the ERPs elicited by pairs with 55 and 65 ms ISI. Although the effect size is bigger in the active than in the passive condition, the statistical comparison between the coefficients was not significant ( $p = .863$  and  $p = .442$  comparing the correlations with the mean and maximum values respectively; the applied function can be found in the Supplementary Material, Section D).

## 4. Discussion

The aim of the present study was to identify electrophysiological markers of the discrimination threshold in auditory temporal processing. Tone pairs were presented with parametrically varying within-pair ISI and their ERP responses were analyzed as a function of ISI in relation to the spatial TOJ threshold obtained in the same participants. In line with our previous findings (Simon et al., 2019), nonlinear amplitude changes were identified, but surprisingly it was significant only in the ‘high-threshold’ group. Similarly to our first experiment of the above cited study, the pronounced change was between the ERPs of tone pairs with  $ISI = 55$  and  $ISI = 65$  ms at the latency range of the P2 component. The latency of the relevant change is in line with previous results showing P2 attenuation with long ISI (Lewandowska, Bekisz, Szymaszek, Wróbel, & Szelag, 2008).

P2 attenuation can be explained by various processes. As in this experiment the attenuation was present even in the passive condition, the hypothesis of a more elaborate conscious processing of long-ISI tones



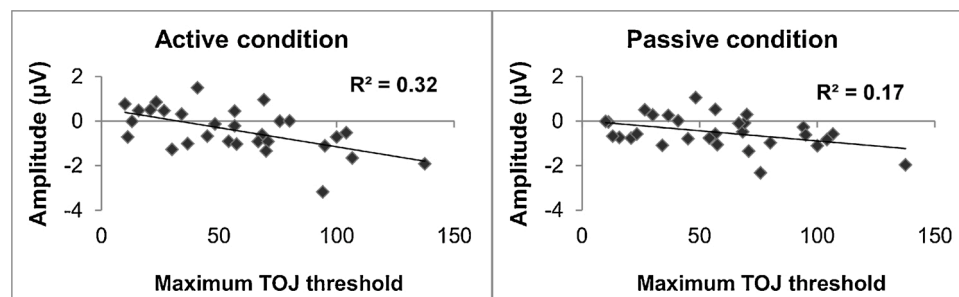
**Fig. 4.** A, ERP amplitudes ( $N = 30$ ) from the early (150-170 ms) measurement window as a function of ISI (x-axis), separately for the active and the passive condition (passive: grey; active: black; collapsed across the two groups with different starting condition as well as the two TOJ threshold groups due to the lack of significant differences). B, Group mean ( $N = 15$ ) average ERP amplitudes from the late (210-230 ms) measurement window as a function of the ISI (x-axis), separately for the active and the passive condition (passive: grey; active: black) and TOJ-group (low TOJ threshold: left panel; high TOJ threshold: right panel; collapsed over the two groups with different starting condition). C, In order to highlight the differences between the adjacent ISI ERPs in the late (210-230 ms) window, the difference amplitudes are presented as a function of ISI pair (x-axis), separately for the two TOJ groups (low threshold: grey; high threshold: black; collapsed across the active and the passive condition and the two groups with different starting condition). \* $p < .05$ , \*\* $p < .01$ .

due to less ambiguity (Lewandowska et al., 2008) is less likely. A higher P2 in the short ISI conditions can be a result of fusion making the consecutive tones one binaural event that produces larger P2 responses than monaural ones (Papesh, Billings, & Baltzell, 2015). Alternatively, the second sound can act like a masking noise that attenuates P2 (Billings, McMillan, Penman, & Gille, 2013). However, we prefer a second N1 explanation, namely that the P2 is attenuated because it co-occurs with the generation of a second N1. This explanation is in line with the visual examination of the scalp distributions but for a definite conclusion more data is needed with more sophisticated procedures (e.g. proper source localization). However, all these explanations are in line with a qualitative change in neural processing as the two tones start to be processed as two distinct auditory events. Another argument in support for the fused nature of the tones at short ISIs is that the N1 did not show its typical attenuating behavior as the ISI decreases (Javitt, Jayachandra, Lindsley, Specht, & Schroeder, 2000); if anything, it was largest at the shortest ISI. Due to the various potential overlap of neural components, it is hard to form a strong conclusion, but these results suggest that while the N1 is an obligatory component, the P2 is a result of a higher level integration process, as no pronounced second P2 emerged.

On the one hand, the results of this study support the claim that with the increase of ISI the differences between adjacent ERP amplitudes (e.g. the ERP of tone pairs having an ISI = 45 ms is adjacent to the ERP of tone pairs having an ISI = 55 ms) are not just the function of the physical parameters but reflect the participants' individual sound processing mechanisms. On the other hand, the appearance of the qualitative change between the ERPs of tones having 55 and 65 ms ISI is not always present or at least identifiable. In our previous study, we considered the change around 60 ms as an event boundary, under which the two sounds are processed as a single event, meaning that the input from the individual tones do not have comparable temporal tags. In this conceptual framework, a lack of pronounced change in the 'low-threshold' group can be explained by either a more distributed threshold blurring a single peak, or with changing thresholds within participant during the recording (e.g. 35 ms at the beginning, 55 ms at the end of the session). However, this conceptual framework is yet to be supported with additional results.

At this point, we can only state that the behavioral spatial TOJ threshold is not just the result of applying different behavioral strategies, but is related to differences in brain processing, which was indicated by the amplitude difference between ERPs of ISI = 55 ms and ISI = 65 ms correlating with the TOJ thresholds. However, if we accept that the boundary around 60 ms does indicate a discrimination threshold, that is in line with the assumed duration of a perceptual moment around 20–70 ms (Fostick & Babkoff, 2013; Pöppel, 1997), then the behavioral thresholds around 20 ms are the result of some participants' ability to use holistic strategies even in the spatial TOJ task.

It should also be noted that the participants of the study of Simon et al. (2019) did not have worse spatial TOJ thresholds than participants



**Fig. 5.** Mean ( $N = 30$ ) ERP amplitude difference in the late (210–230 ms) measurement window between the 55 and 65 ms ISI (y-axis) as a function of the TOJ threshold (x-axis) with the active condition shown on the left, the passive on the right panel. A larger negative amplitude difference reflects larger absolute difference between the ERP amplitudes. Trend lines were fitted and their respective  $R^2$  values were added for easier comparison.

in the current study, yet the amplitude differences showed a robust nonlinearity (Simon et al., 2019). This might indicate the context dependency of tone processing at the examined time scale. In both experiments, they started with the behavioral measurements, but the later EEG recording could have been a different experience for the two groups of the two studies. One source of difference between the two contexts was that pauses were introduced in the current experiment. The context dependency of sound processing at this sub-100 ms time scale is in line with the relatively low test-retest reliability of the spatial TOJ task (Fink et al., 2005) or its susceptibility to deterioration as a function of measurement repetition (Simon et al., 2020).

Temporal processing of quick acoustic changes is a hot topic in language processing as temporal discrimination of auditory events of a few tens of milliseconds were proposed to be suboptimal in developmental dyslexia (Farmer & Klein, 1995; Fostick, Eshcoly et al., 2014; Gaab et al., 2007). One popular explanation of developmental dyslexia and specific language learning impairment is the rapid temporal processing deficiency hypothesis (2004, Ben-Artzi, Fostick, & Babkoff, 2005; Tallal, 1980). However, the theory is debated as the supporting evidence is contradictory (Banai & Ahissar, 2004; Bishop, Carlyon, Deeks, & Bishop, 1999; Goswami, 2011; Ramus et al., 2003; Ziegler, Pech-Georgel, George, & Lorenzi, 2009). Besides the heterogeneous nature of the condition, one of the reasons of the contradicting measures can be that behavioral tests claimed to assess temporal processing measure different constructs and there is no clear mapping between the behavioral index and the underlying constructs (Protopapas, 2014). Therefore, further exploring the electrophysiological changes associated with these paradigms proposed to investigate temporal discrimination can be beneficial in understanding the exact nature of the different temporal processing functions.

## 5. Conclusion

In this study, we showed that event-related responses of tone pairs separated by short (25–75 ms) time intervals do not only result from physical differences between stimuli but are also related to the spatial TOJ threshold of the individuals. Furthermore, applying specific strategies may introduce a bias on behavioral sTOJ thresholds, as the behavioral results of the best performing group were much better than the threshold indicated by the pronounced change in the ERP amplitudes.

## Declaration of Competing interest

We declare no conflict of interest.

## Acknowledgement

We are grateful for Dr. István Winkler, who made the conduction of this experiment possible and who contributed to the manuscript with

many useful comments.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.biopsycho.2020.107947>.

## References

- Banai, K., & Ahissar, M. (2004). Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiology and Neurotology*, 9(6), 328–340.
- Ben-Artzi, E., Fostick, L., & Babkoff, H. (2005). Deficits in temporal-order judgments in dyslexia: Evidence from diotic stimuli differing spectrally and from dichotic stimuli differing only by perceived location. *Neuropsychologia*, 43(5), 714–723.
- Bernasconi, F., Manuel, A. L., Murray, M. M., & Spierer, L. (2011). Pre-stimulus beta oscillations within left posterior sylvian regions impact auditory temporal order judgment accuracy. *International Journal of Psychophysiology*, 79(2), 244–248.
- Bernasconi, F., Grivel, J., Murray, M. M., & Spierer, L. (2010a). Interhemispheric coupling between the posterior sylvian regions impacts successful auditory temporal order judgment. *Neuropsychologia*, 48(9), 2579–2585.
- Bernasconi, F., Grivel, J., Murray, M. M., & Spierer, L. (2010b). Plastic brain mechanisms for attaining auditory temporal order judgment proficiency. *Neuroimage*, 50(3), 1271–1279.
- Billings, C. J., McMillan, G. P., Penman, T. M., & Gille, S. M. (2013). Predicting perception in noise using cortical auditory evoked potentials. *Journal of the Association for Research in Otolaryngology*, 14(6), 891–903.
- Bishop, D. V. M., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: Neither necessary nor sufficient for causing language impairment in children. *Journal of Speech Language and Hearing Research*, 42(6), 1295–1310.
- Cutting, J. E. (1976). Auditory and linguistic processes in speech perception: Inferences from six fusions in dichotic listening. *Psychological Review*, 83(2), 114.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, 134, 9–21.
- Farmer, M. E., & Klein, R. M. (1995). The evidence for a temporal processing deficit linked to dyslexia: A review. *Psychonomic Bulletin & Review*, 2(4), 460–493.
- Fink, M., Churan, J., & Wittmann, M. (2005). Assessment of auditory temporal-order thresholds—A comparison of different measurement procedures and the influences of age and gender. *Restorative Neurology and Neuroscience*, 23(5, 6), 281–296.
- Fostick, L., & Babkoff, H. (2013). Different response patterns between auditory spectral and spatial temporal order judgment (TOJ). *Experimental Psychology*.
- Fostick, L., Eshcoly, R., Shtibelman, H., Nehemia, R., & Levi, H. (2014). Efficacy of temporal processing training to improve phonological awareness among dyslexic and normal reading students. *Journal of Experimental Psychology Human Perception and Performance*, 40(5), 1799.
- Fostick, L., Wechsler, S., & Peretz, E. (2014). Short-term learning effect in different psychoacoustic measures. *Journal of Basic and Clinical Physiology and Pharmacology*, 25(3), 307–312.
- Gaab, N., Gabrieli, J. D. E., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology and Neuroscience*, 25(3–4), 295–310.
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3–10.
- Halliday, L. F., Tuomaine, O., & Rosen, S. (2017). Auditory processing deficits are sometimes necessary and sometimes sufficient for language difficulties in children: Evidence from mild to moderate sensorineural hearing loss. *Cognition*, 166, 139–151.
- Javitt, D. C., Jayachandran, M., Lindsley, R. W., Specht, C. M., & Schroeder, C. E. (2000). Schizophrenia-like deficits in auditory P1 and N1 refractoriness induced by the psychomimetic agent phencyclidine (PCP). *Clinical Neurophysiology*, 111(5), 833–836.



- Kanabus, M., Szélag, E., Rojek, E., & Pöppel, E. (2002). Temporal order judgement for auditory and visual stimuli. *Acta Neurobiologiae Experimentalis*, 62(4), 263–270.
- Lewandowska, M., Bekisz, M., Szymaszek, A., Wróbel, A., & Szélag, E. (2008). Towards electrophysiological correlates of auditory perception of temporal order. *Neuroscience Letters*, 437(2), 139–143.
- Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213.
- May, P. J. C., & Tiitinen, H. (2007). The role of adaptation-based memory in auditory cortex. In *International Congress Series* (Vol. 1300, pp. 53–56). Elsevier. doi: 10.1016/j.ics.2007.01.051.
- Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48(2), 229–240.
- Papesh, M. A., Billings, C. J., & Baltzell, L. S. (2015). Background noise can enhance cortical auditory evoked potentials under certain conditions. *Clinical Neurophysiology*, 126(7), 1319–1330.
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences*, 1(2), 56–61.
- Protopapas, A. (2014). From temporal processing to developmental language disorders: Mind the gap. *Philosophical Transactions Biological Sciences*, 369(1634), Article 20130090.
- Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., ... Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126(4), 841–865.
- Saija, J. D., Başkent, D., Andringa, T. C., & Akyürek, E. G. (2019). Visual and auditory temporal integration in healthy younger and older adults. *Psychological Research*, 83(5), 951–967.
- Simon, J., Balla, V., & Winkler, I. (2019). Temporal boundary of auditory event formation: An electrophysiological marker. *International Journal of Psychophysiology*, 140, 53–61.
- Simon, J., Takács, E., Orosz, G., Berki, B., & Winkler, I. (2020). Short-term cognitive fatigue effect on auditory temporal order judgments. *Experimental Brain Research*, 1–15.
- Steinschneider, M., Volkov, I. O., Fishman, Y. I., Oya, H., Arezzo, J. C., & Howard, M. A., III (2004). Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cerebral Cortex*, 15(2), 170–186.
- Steinschneider, M., Fishman, Y. I., & Arezzo, J. C. (2003). Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *The Journal of the Acoustical Society of America*, 114(1), 307–321.
- Stevenson, R. A., Park, S., Cochran, C., McIntosh, L. G., Noel, J. P., Barense, M. D., ... Wallace, M. T. (2017). The associations between multisensory temporal processing and symptoms of schizophrenia. *Schizophrenia Research*, 179, 97–103.
- Szélag, E., Lewandowska, M., Wolak, T., Seniow, J., Poniatowska, R., Pöppel, E., et al. (2014). Training in rapid auditory processing ameliorates auditory comprehension in aphasic patients: A randomized controlled pilot study. *Journal of the Neurological Sciences*, 338(1–2), 77–86.
- Szymaszek, A., Sereda, M., Pöppel, E., & Szélag, E. (2009). Individual differences in the perception of temporal order: The effect of age and cognition. *Cognitive Neuropsychology*, 26(2), 135–147.
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5(9), 721.
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, 9(2), 182–198.
- Zaehle, T., Jancke, L., & Meyer, M. (2007). Electrical brain imaging evidences left auditory cortex involvement in speech and non-speech discrimination based on temporal features. *Behavioral and Brain Functions*, 3(1), 63.
- Zaehle, T., Wüstenberg, T., Meyer, M., & Jäncke, L. (2004). Evidence for rapid auditory perception as the foundation of speech processing: A sparse temporal sampling fMRI study. *The European Journal of Neuroscience*, 20(9), 2447–2456.
- Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12(5), 732–745.